Collective Copyright

Enabling the Natural Evolution of Content Creation in the Web Era

http://collectivecopyright.com

Emanuele Lunadei
Helikonia Indi SRL
lele.lunadei@helikonia.it

Christian Valdivia Torres
Accenture Italy
c.torres@accenture.com

Erik Cambria
MIT Media Laboratory
cambria@media.mit.edu

ABSTRACT

The way people create and share content has radically changed in the past few years thanks to the advent of social networks, web communities, blogs, wikis, and other online collaborative media. Such online social data are continuously growing in a way that makes it difficult to efficiently aggregate them, since they are the expression of a multitude of single content creators that most of the times show only a small percentage of originality. The act of 'sharing' is still tied to a pre-Internet fashion that sees it as a step following (and never preceding) content creation, as enforced by the rules of publishing and copyright. In the Internet era, the pieces of the puzzle of a valuable work might be scattered throughout the whole Web. In order to hinder the obsolete create-then-share trend that is killing creativity and usefulness of the Web, we propose a new copyright framework, which allows content to be shared while being created, in a way that this can gain increasing value as it becomes part of an increasingly richer puzzle.

Categories and Subject Descriptors

D.2.9 [**Management**]: Copyrights;
H.3.5 [**Online Information Services**]: Data sharing

Keywords

Collective Copyright, content creation, content sharing, collective intelligence, wisdom of crowds

1. INTRODUCTION

The world is constantly changing, and so are people, and the Web. With the dawn of the Internet Age, civilization has undergone profound, rapid-fire changes which we are experiencing more than ever today. Even technologies that are adapting, growing, and innovating give us the feeling that obsolescence is right around the corner. Such a fast-evolving environment requires a continuous emergence of technological and social changes.

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2580071.

In the current era of on-line content sharing and consumption, for example, the concept of copyright no longer applies to the ongoing dynamics of the World Wide Web. On the distribution end, it is almost impossible to control both copyright infringement and the sharing of copyrighted material, on the production end it feeds individuality producing an overload of amateur content (which is why, nowadays, the 'wisdom of crowds' still corresponds to mere *collected*, rather than *collective*, intelligence [8]).

To this end, we propose Collective Copyright, a novel framework that allows Web users to effectively share and enhance their content during creation through a certified framework. Collective Copyright is a modernization of the traditional concept of copyright, enabled by already existing dynamics and technologies, that allows users to enhance the value of their content as if they were *one single creator*. The current feeling associated with the concept of sharing content on the Web is often fear of losing the intellectual property these generate. Through Collective Copyright, such fear will eventually turn into eagerness to receive feedbacks and contributions from the Web community, or even let others finish what one has started as far as the enhancement and evolution of the shared content will be done in a trusted environment, hence, providing the means for true *collective* intelligence. For this reason, we adopt Conway's *glider* [5] as the logo of Collective Copyright (Figure 1), meaning that simple rules can trigger infinite chain reactions.

In this work, we apply Collective Copyright to the world of music production, but we believe that the framework can be extended to any kind of content that can potentially be created online. The rest of this paper is organized as follows: the first section presents the evolution of music production from the era of *Sheet Music*, through the age of *Recorded Music* to the next era of *Collective Music*; the following section describes related work in the field of content co-creation and shared copyright; the next two sections introduce and illustrate the architecture of the proposed framework; after a proof of concept, finally, some concluding remarks and future work recommendations are made.

2. BEYOND RECORDED MUSIC

Today, quite a few old-fashioned record producers still like to claim they are searching for the *New Beatles*. So far, however, nobody has found them yet. Clearly, such producers do not refer to an exact copy of the Fab Four, but rather to a product that has winning features such as longevity in the market and balance between commercial and niche quality product.

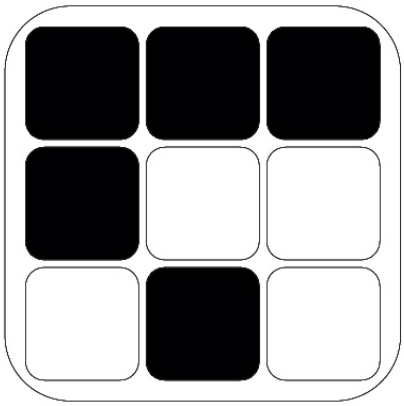

Figure 1: Collective Copyright logo

In today's fast-evolving market, the consumption of music (and content in general) is migrating from the concept of ownership to the concept of everlasting online availability, so that old-fashioned producers have to resign to the assumption that the record industry is in bankruptcy because records are no longer commodities. We have to keep in mind that we are living in a particular time window of the history of music that followed the age of *Sheet Music*. Today, we live in the *Recorded Music* era, which saw its peak from 1965 to 1995 and has created a work of art all to itself, which was meant to produce revenue: the *Record*. Despite the common belief, such an age is no longer contemporary and has to evolve into a new era: the age of *Collective Music*.

Comparing the past era to the present one would be like thinking of Beethoven writing the 9th Symphony for a particular conductor and orchestra with the intention of making a record out of it and the concept that everyone else playing such work would be "covering the 9th Symphony". The uniqueness of performer (artist, act, interpreter, songwriter) and performance (song, music, content in general) is a concept that belongs only to the *Recorded Music* age. The opposite to the previous example would be The Police writing sheet music of "Every breath you take" without recording it: it would be pointless since the features of the original recorded track are part of the content itself.

In the passage from a model built on constant creation and performance to a model of recording and collecting royalty revenues, we have lived a golden age of economic stability that was based on scarcity of content and no possibility of free sharing. The artists and contents themselves have been an outcome of such model which is no longer sustainable in an era of abundance and free sharing, and will no more produce such artists and works.

As history teaches us, mankind's universal works of art are not a random manifestation of creativity, but rather the product of the features and socio-economic dynamics of humanity in a specific time window. It would not be fair to state that today nobody is skilled enough to be a painter such as one of the Baroque period. Instead, it could be said that the present socio-economic conditions are statistically not likely to produce a Caravaggio. The analysis of the intrinsic features of humanity in this historical period and the virtuous mechanisms potentially leverageable is a viable way to the future creation of universal works of art.

In the context of music production, this means evolving from the era of *Recorded Music* to the age of *Collective Music*, where the performance results from the *collective creativity* of a multitude of performers and social media interactions regulated by the Collective Copyright.

2.1 Who Is Going to Pay for Music?

Although our goal is not how to find new ways to sell music but rather how to produce music worth paying for, it is relevant to focus on the situation as of 2014. Music streaming services such as Spotify[1] are replacing retail stores not only in the real world but also in the digital one. Subscription-based streaming is taking over legal (or illegal) download as long as the service is convenient and accessible [14]. We are facing a new revenue model in the music business, based on the Open Music Model system [6], that works on collecting subscription fees from users and then compensating artists according to their streaming rate.

In this scenario a Peter Gabriel fan pays 10 dollar a month to listen to music as well as Peter Gabriel himself. The question to be answered is no longer "Who is going to buy a song?" but "Who is going to listen to it?" since everyone is paying the same 'ticket' for access. As we will discuss later, our framework allows for the creation of multiple versions of a single content, but we believe that, by virtue of natural selection, only one version will gain traction: the one that will be best written and produced.

3. RELATED WORK

The tension to a modernization of the copyright law can be found in widely used models such as Copyleft, Open Source and Creative Commons. The key motivation for the creation of such frameworks is to allow the voluntary sharing of content that would otherwise be prevented by the very nature of copyright. Besides the special case of Linux O.S., a tangible experience of collaborative creation in computer science, the above-mentioned frameworks have not produced a trust infrastructure and a sharing model of the potential revenues sufficient to trigger a real process of collaborative content creation on a large scale in other creative fields.

Creative Commons[2] in most cases is used as a license to distribute user-generated content (UGC) in non-commercial mode, which is conceptually identical to inserting one's work voluntarily in the public domain [13]. Furthermore, the remix culture is only partly a creative process, the existing platforms collecting audio samples provided by users to be remixed do not have the potential to trigger a collaborative creation process that would require to work in real time and not by assembling samples no longer editable. More clearly, the remix culture will only apparently reproduce the collaborative process that lays at the bottom of the music production.

3.1 Music Production in Education

An important project that brought us to the idea that we are presenting was developed from 2005 to 2013 with high school students in a recording studio built by the municipal authority of Rome[3].

[1] `http://spotify.com`
[2] `http://creativecommons.org`
[3] `http://comune.roma.it/wps/portal/pcr?jppagecode=prog_rmrock_rmpop_dses.wp`

The facility provided free access for hundreds of musicians aged 15-19 coming from public high schools. We decided not to use it as a place for teaching music related subjects but rather for producing original material written by the students in an old school fashion, as if we were in a record label. A Record producer was hired and we invited young Sound engineers from Music Academies to do internships, as well as young session musicians.

The best skills were leveraged, some wrote music, some wrote lyrics (often the best lyricists were not musicians at all). We tested a collaborative framework, very similar to a beta version of Collective Copyright, which was widely accepted, creating a non-stop workshop that crossed geographical barriers also with other cities. The Record production has identified the user's potential to produce content at a very good level when placed in a collaborative dynamic, furthermore this age group is still creatively 'untouched' by the outside world and capable of writing music which is far from mainstream[4]. Over 90% showed openness to let third parties complete and/or use their content when granted the right share of intellectual property. We believe that such projects, developed as co-curricular activities at first, can be the laboratories to test and spread the mindset to this framework.

4. A NEW COPYRIGHT FRAMEWORK

The Web is continuously evolving and Web users are evolving with it. With the evolution of the Web from read-only to read-write, the Web user evolved to the stage of *prosumer* – producer and consumer at the same time. Thanks to new affordable tools that enable them to produce audio and video in a semi-professional fashion, today prosumers can run all alone the production/distribution and marketing phase of their UGCs.

In the current situation, where mainstream mechanisms have poured into commercial dynamics of entertainment, we find ourselves in a stand-by phase, within a system that is no longer economically sustainable, waiting for the independent network to produce new records comparable to "Sgt Pepper".

The need for a new framework for co-creation of UGCs, at least in the music industry, is determined by the following factors:

1. like in the Caravaggio example, classic discography of the period 1965-1995 is the socio-economic outcome of music publishing, copyright and business of non-duplicable physical record at their historical peak. It is a thirty-year timespan that, in an overhead view, should be taken as an exception and not as a rule.

2. the work of art is the codification of collective elements catalyzed by a non-ordinary singularity and assisted through production by an economically stable industry.

3. a hyper-fragmentation of the valuable creativity is not sustainable, even assuming that every UGC is a likewise product this would create an inability to fruition given by overload of contents.

[4]`http://itunes.apple.com/us/album/voci-della-periferia/id511634917`

4.1 The Puzzle

To use an example, if we took the shuffled pieces of a hundred puzzles of 1000 pieces each with the intent to put together only one puzzle displaying a known image that we define "valuable", with the current technology we would be able to reduce exponentially the operation, manually achievable in the longer term. With a small upgrade, we could have the possibility to reconstruct all of the hundred puzzles, in this case the operation would be manually possible but way longer. With a further upgrade, we could aim to put together all the puzzles without knowing the displayed images, assuming they are not senseless but have a balance. In this case, however, the operation done manually would be so long as to dissuade the will to achieve it [17].

If we consider the potential that in the hundred puzzles there is a small percentage of valuable works, we can state that with the current technology it makes sense to do a search [8]. We use this metaphor to suggest that in history all the innovations that have revolutionized the market by destroying an old model can be used to reconstruct it in a new key [3, 4, 11].

4.2 Framework Modules

The technologies necessary to build a new effective framework for content co-creation and co-sharing are already available, but they need to be put together in a more meaningful and content-centric way. Such technologies are the following:

Acoustic Fingerprint
Currently used by softwares such as Shazam, to look for matches between an audio file being played and the ones contained in a database, or in the Content ID that keeps a log of YouTube content, using a service that automatically checks if there is copyright-protected material in the uploaded video. In Collective Copyright, a similar technology [16] is used for mapping and cataloging demos by building up a revision history (on the Wikipedia model) that replaces the classic deposit via musical score.

Music Genre Classification
Music classification algorithms are used by music recommendation platforms such as Pandora for classifying music files according to a variety of specifications. Collective Copyright exploits a recent technology developed by Poria et al. [15] as a first subdivision of potentially matchable audio files, not aiming to mix them together but rather to suggest a link among users who have worked on similar projects.

Wikipedia-Like Revision History
Collective Copyright allows authors the possibility to open different branches of the same demo, hence enabling multiple versions of the same work to be created. Interactions in the community of users (creators or non-creators) will trigger natural selection of the best solutions.

Rating/Reputation
This mechanism was crucial for the trust consolidation in widespread platforms such as eBay. In Collective Copyright, it performs the following functions:

1. first step visibility of the content recommendation and semi-finished products to potential connections between creative users

2. natural selection of choices during the creative phase through interaction with the community [12].

3. internal balance through compliance with the policy (e.g., stable sellers on eBay as well as the YouTube channel with constant production that benefit from the monetization perfectly meet the platform's policies not to be excluded from it, and resetting the rating, history, and reputation, which are tangible elements that are built over time).

4. natural selection of users, both in the composition and the production phase, bringing back the figure of the session man, the high skilled musician that used to work for recording studios. Actually, all the figures of the music production chain can be embedded within the framework[5].

Natural Language Processing
Assist the creation and search phase of the right solutions in the music lyrics, assist the search for the right user profile for a specific collaboration, analysis of data text to guide possible connections between users or between projects in progress through the use of novel semantics-based natural language processing techniques for analogical reasoning [1] and crowd validation [2].

YouTube Video Response Dynamics
When a YouTube user brings an added value to a given topic by uploading it as a video response he's actually publically documenting his research contribution, which often becomes part of the final result.

4.3 Copyright Sharing Example

We hereby propose an example of content co-creation regulated by Collective Copyright, which shows how intellectual property is dynamically shared while content is modified (Figure 2). In this example, the shares are standardized but will be regulated by an internal policy that will take into account a series of variables, depending on the quantity of contribution and at which step the contribution takes place.

User A (composer) starts the development of content *x* (e.g., a song/music piece). So far he owns 100% of the share, which we term content *x.1*. User B makes the first change to content *x*, taking it to revision *x.2*. At this time, both users hold a 50% share of *x.2*. The development hereinafter involves users C and D, bringing the content to *x.3* and *x.4*. If the process ends up here, developers A, B , C and D would hold 25% of the share (or 'royalties', as defined in the music market) of *x.4* each.

At this point, user E steps in and brings the content to revision *x.5*, users A, B, C, D and E are satisfied by the result and decide to bring it to production phase. User F decides to make a different change in the phase *x.4* (e.g., lyrics), a different revision from the one proposed by user E and not accepted by the first four developers, nor by user E. User F is free to do so: the system tags its modification as *x.5.1* and opens on a second branch of the content under development. Content *x.5.1* may be published, more likely by user F along with subsequent developers who carry out its variant.

[5]http://firstmonday.org/ojs/index.php/fm/article/view/1975/1850

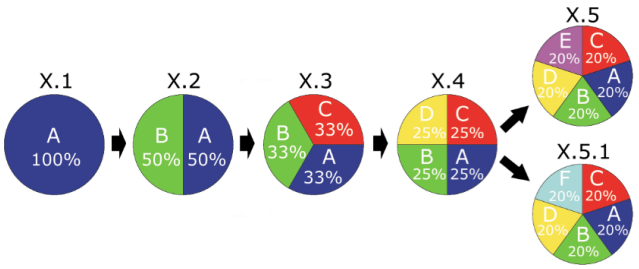

Figure 2: Copyright sharing during content co-creation

The key concept here is that, after revision *x.5.1* takes place, users A, B, C, D and F hold 20% each of *x.5.1*, as well as A, B, C, D and E hold 20% each of *x.5*. The user has to accept subsequent changes to his revision as long as his share is respected. Any user who does not approve a subsequent change is free to stick to the previous version and bring it to production, he will still hold the same share of the content developed until that point.

Each demo is:

1. registered with a unique ID (through Acoustic Fingerprint)

2. catalogued (by means of Music Genre Classification)

3. public and available for both developers and the community

4. editable respecting the developers percentages as in the above example

5. rated and commented

In this context, the classic concepts of *published* versus *unpublished* lose value in favor of *consolidated* versus *non-consolidated* plus *highly rated/viewed* versus *low rated/viewed*. Not only authors take part of the share of content *x.n*, but also those who participate in the production phase that brings a *demo* to be the final song.

5. FRAMEWORK ARCHITECTURE

The Collective Copyright framework will be available as an online platform to be used both as a stand-alone tool and a social-networking framework (e.g., Instagram) with a link button on related platforms such as YouTube and SoundCloud (Figure 3). The effectiveness of the framework is the result of two parallel sub-systems, namely the *Collective Content Development System* and the *Trust-Based Copyright System*.

The former consists of mechanisms of interaction to aid the search for added value (e.g., networking/crowdsourcing, reputation/rating, representation of under development content, etc.) and to efficiently spread higher added value content (e.g., music recommendation engine, circles of people with similar tastes, etc.). The *Trust-Based Copyright System*, instead, consists of an eBay-like reputation mechanism so that it creates much more value to users that respect the policies and build a reputation than to those who try hacking the rules and risk being put out of it.

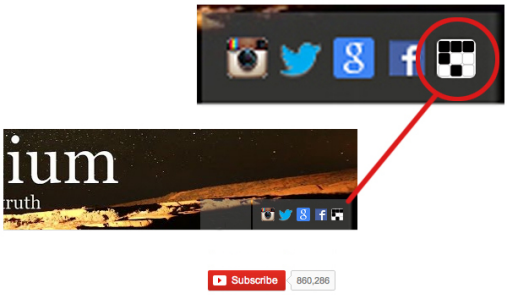

Figure 3: Web users will be able to share their content with a simple click, in the same wake as they currently access Facebook and its like

The subsystem, moreover, ensures that everyone can have multiple shares of numerous contents. Finally, the subsystem holds a fixed percentage (e.g., 10%) on each production, in order to avoid having a revenue model built upon ads.

5.1 Collective Content Development System

The Collective Content Development System handles mechanisms of value creation and, hence, takes care of enabling, triggering, and leveraging creativity through the platform. It models user interaction (both as consumers and creators of music) through the following stages:

Content/Contact Notification/Recommendation
This stage is efficient if fully integrated with social networks. Besides authors and musicians, an important role is played by the community (e.g., the right lyricist can be found through users suggestion). The engagement of blogs users is an example of often wasted data due to non structured interactions [10].

Content Sharing
Content can be uploaded in the Collective Copyright framework or linked from a different UGC platform, e.g., YouTube and SoundCloud. In the latter case, the framework will only handle cataloging functions and link dynamics. Users will also be able to share audio via social-networking apps, e.g., WhatsApp[6], and digital audio workstations, e.g., Ohm Studio[7], which will become increasingly effective and adopted, as soon as the broadband will allow it.

Content Evolution Visualization
The graphical approach to branching (geographical, social, etc.) the creation-evolution of contents will be a work of art itself, in the same wake as in Worldmapper[8] territories are re-sized on each map according to the subject of interest.

Engagement Attraction
In recording studios, the lack of quality of still unpublished demos is often compensated by the authors' attraction to it, due to their production engagement. This entices them to listen to demos over and over again, even though they are merely rough versions of the future track. This vision is not shared by outside listeners, who are only interested in the final version. In a scenario where groups of users spend time listening to rough versions, this trend is key to measure the potential of a demo while it is reaching its final stage.

[6]`http://www.whatsapp.com`
[7]`http://www.ohmstudio.com`
[8]`http://www.worldmapper.org`

5.2 Trust-Based Copyright System

Similarly to the Semantic Web stack [7], the Collective Copyright framework builds upon the concept of trust. User reputation is comparable to the one of eBay sellers, which allows them to become more trustable as their contribution to the platform and respect of the policies grows over time.

In the case of composers, they will grow in ratings and gain more chances to collaborate on valuable projects. In the case of musicians, they will grow in visibility and will be required to play in different tracks on a global scale, in same wake as it used to be done for session musicians working for record labels. A user with consolidated reputation will do his/her best to avoid losing his *Collective Copyright account*, since he/she would have to start all over again. The *Trust-Based Copyright System* also ensures the application of basic terms and conditions of the framework, namely:

1. The framework is a collaborative platform for content development. Its purpose is to enable and optimize the mechanisms of interaction between users connected via social media, with the aim of creating content with higher added value.

2. Any content developed within the platform is protected by the Collective Copyright policy. The use of the material produced accepts compliance with the rules of the platform. Infringement of one of them leads to exclusion from the platform and deletion of the account and the history connected to it.

3. The mapping of the steps in the development of contents is proof of deposit of intellectual work in accordance with current rules that govern it. The platform is in charge of mapping and sharing and not of legal assistance in case of controversy.

4. Users are aware that their contributions can be consulted at any stage of development.

6. PROOF OF CONCEPT

Although we believe that the Collective Copyright framework applies to all creative fields, this paper focuses on the Record industry, in which the production process behind it provides a mindset and a set of practices that can be easily converted into the new model without making it too articulated.

Collaborative Production Process
The entire catalog of the world discography is the result of a collaboration during all stages of production, passing through the evaluation of an external point of view given by the producer and so on all the way to market launch.

Outsourcing Mindset
In songwriting, the Lyricist is often considered as a co-author. Historical examples of such partenrships between composer and lyricist can be seen in the example of Elton John and Bernie Taupin [9].

Share Split A Priori
The percentage share of the collaborators on the final product shall be decided a priori (e.g., 50% each for two authors or 33% in the case of three), although later the individual contributions will be variable. This model is widely accepted and is perfectly applicable to the policy of a system to sign a priori imposing a standardized division of the share.

Remotely Working Dynamics
In a recording studio, the control and recording rooms are acoustically isolated and communications are done via a talkback microphone. The visual contact between recording musicians and people in the control room is through a small window and more often a webcam. With the right mindset, users could achieve the same result through a Skype-like environment, as shown by already existing remotely collaborative audio workstations such as Ohm Studio.

7. CONCLUSION

In the era of democratic connection among users on the Web, it is inevitable to evolve from a model of creators of single contents, ratio 1:1 (imposed by the classic copyright mindset) to a model of various creators of single contents, ratio n:1. To this end, we propose a deep modernization of the concept of copyright, termed Collective Copyright, which exploits already existing dynamics and technologies to allow users to share their content in a way that this can gain increasing value.

The current prototype is applied to the context of music production, but Collective Copyright can be potentially extended to any kind of content that is created and shared online. Collective Copyright is currently being pilot tested by a small group of young musicians. Soon, we will employ the same user interface prototype (Figure **??**) in the context of screenplay production.

The current prosumer mindset is a transient phase where the possibility of free sharing is interpreted as the ultimate goal of the creative process, producing an overload of contents. A new model is necessary at this point that leverages the potential of individual creators within the creative phase and not after publication. It is actually only an explicit view of the collective intelligence manifestation that in the past was implicit in art, culture, and science.

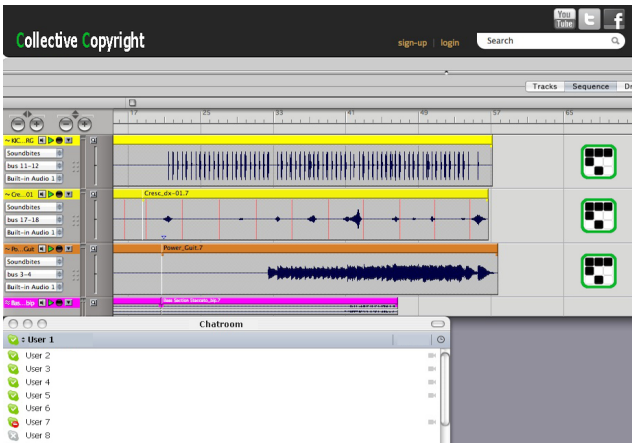

Figure 4: Collective Copyright user interface

The new Lennon/McCartneys will not be only two people, but a network of co-authors and user interactions that will replace the entire Record industry by re-establishing an equilibrium between the amount of produced contents, the value of them and the possibility of mass consumption.

If we were to define what is the 'mass virtuosity' of the current era, it would probably be *interaction*: the ability to abstract and imagine oneself as a single part of a system. What is lacking today is an economic model that triggers and promotes such dynamics.

8. REFERENCES

[1] E. Cambria, P. Gastaldo, F. Bisio, and R. Zunino. An ELM-based model for affective analogical reasoning. *Neurocomputing*, Special Issue on Extreme Learning Machines, 2014.

[2] E. Cambria, A. Hussain, C. Havasi, C. Eckl, and J. Munro. Towards crowd validation of the UK national health service. In *WebSci*, Raleigh, 2010.

[3] E. Cambria and B. White. Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, 9(2), 2014.

[4] D. Engelbart. Boosting our collective IQ. *Communications of the ACM*, 38(8):30–33, 1995.

[5] M. Gardner. Mathematical games – The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223:120–123, 1970.

[6] S. Ghosemajumder. Advanced peer-based technology business models. Master's thesis, MIT Sloan School of Management, 2002.

[7] J. Golbeck, B. Parsia, and J. Hendler. *Trust networks on the semantic web*. Springer, 2003.

[8] T. Gruber. Collective knowledge systems: Where the social web meets the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1):4–13, 12 2007.

[9] R. Hass, R. Weisberg, and J. Choi. Quantitative case-studies in musical composition: The development of creativity in popular-songwriting teams. Technical report, Temple University, 2010.

[10] J. Hendler. Web 3.0 emerging. *IEEE Computer 42*, 42(1), 2009.

[11] M. McLuhan. *Understanding Media: The Extensions of Man*. McGraw-Hill, 1964.

[12] S. Nambisan. Platforms for collaboration. Technical report, Stanford University, 2009.

[13] M. Nielsen. *Reinventing Discovery: The New Era of Networked Science*. Princeton University Press, 2011.

[14] S. Parker. The daily beast's innovators summit, 2010.

[15] S. Poria, A. Gelbukh, A. Hussain, S. Bandyopadhyay, and N. Howard. Music genre classification: A semi-supervised approach. In *MCPR*, pages 254–263, 2013.

[16] E. Principi, S. Squartini, E. Cambria, and F. Piazza. Acoustic template-matching for automatic emergency state detection: An ELM based algorithm. *Neurocomputing*, Special Issue on Extreme Learning Machines, 2014.

[17] J. Surowiecki. *The Wisdom of Crowds*. Anchor, New York, 2005.